Non-verbal communication through Virtual Humans that can communicate

Daniel Thalmann
EPFL, Computer Graphics Lab
CH 1015 Lausanne, Switzerland
Daniel.Thalmann@epfl.ch
http://ligwww.epfl.ch

The merging of recent developments in Virtual Reality (VR), Human Animation and Video Analysis techniques has led to the integration of Virtual Humans in VR. Interaction with Virtual Environments may be at various level of user configuration. A high end configuration could involve an immersive environment where users would interact by voice, gesture and physiological signals with virtual humans that would help them explore their digital data environment, both locally and over the Web. For this, Virtual Humans should be able to recognize gestures, speech and expressions of the user and answer by speech and animation. The ultimate objective in creating realistic and believable virtual actors is to build intelligent autonomous virtual humans with adaptation, perception and memory. These actors should be able to act freely and emotionally. Ideally, they should be conscious and unpredictable.

Most of today's virtual environments are populated with some kind of autonomous life-like agents. Such agents follow a pre-programmed sequence of behaviours that exclude the user as a participating entity in the virtual society. In order to make inhabited virtual reality an attractive place for information exchange and social interaction, we need to equip the Virtual Humans with some perception and interpretation skills. An important skill is human action recognition. By opposition to human-computer interfaces that focus on speech or hand gestures, we have proposed a full-body integration of the user..

One goal in VR is to simulate real world situations. As a consequence, the interface has to facilitate the real world interaction paradigms. However, reproducing every single detail of an interaction is not necessarily a desirable goal. Grasping an object is such an example. It is relatively easy to copy a performer's hand motions onto virtual hands. Even collision detection between objects and the virtual hands can be mastered [1]. Nevertheless, experience shows that grasping a virtual object is still more difficult than in real life. The problem is the poor quality of the feedback loop. Neither can we manage realistic 3D-vision sensation, nor do we have realistic force feedback devices. We think that grasping an object with detailed finger control is not necessarily desirable. Instead of asking the user to perform a precision grasp, we ask the user to perform a grasp gesture. This action is recognized by the system and the object snaps to his/her hands. Of course, in real life, objects generally do not behave like magnets, but this can be an acceptable compromise between the realism of a task performance and the convenience of the interface. Other interaction tasks that can be advantageously simplified through action recognition events are walking, sitting or simply looking around by turning the head. In most VR systems, such actions are triggered by some hand posture recognition events. We claim that such actions can be more naturally triggered through a full-body action recognition system.

As technology improves, VR interfaces based on body actions and expressions will become more and more important. In the domain of games, an evident application is the control of the hero by body actions. The robotics domain is another area where such an interface presents an attractive issue, especially because telepresence is a hot topic nowadays. We think that we need an action model along with a real-time recognition system and show how it may be used in Virtual Environments. The real people are of course easily aware of the actions of the Virtual Humans through VR tools like Head-mounted displays, but one major problem to solve is to make the virtual actors conscious of the behaviour of the real people. Virtual actors should sense the participants through their virtual sensors. Such a perceptive actor would be independent of each VR representation and he could in the same manner communicate with participants and other perceptive actors. The real time constraints in VR demand fast reaction to sound signals and fast recognition of the semantic it carries. For the interaction between virtual humans and real ones, gesture recognition is a key issue. As an example, Emering et al. [2] describe a system for interaction between a real person represented by an avatar and an autonomous actor (Fig.1). The motion of the real person is captured using magnetic trackers. The gestures are recognised by the system and the information is transmitted to the virtual actor who is able to react to the gestures and decide which attitude to do. An Action Recognition Algorithm recognition process exploits a multi-level action model in order to perform in real-time. First, the Candidate Action Set is initialized with the whole action database. Then, the motion capture system provides the user's posture directly in terms of joint angles. This information drives the action selection process at the five levels of the action model. The ARA retains only those actions which match the current action's characteristic or which do not define any action primitive at that level. The hierarchical approach overcomes the limitation resulting from the extreme simplicity of the matching algorithm while allowing real-time performance. For each new body posture sample the body movement is analyzed to derive the average velocity of the Center of Mass and End Effectors over a short period of time.

But the most promising area is Augmented Reality with real humans interacting with virtual ones in the same scene. In this case, a standard video camera should be used for the real-time gesture recognition, which implies real-time tracking algorithms. Balcisoy et al. [3] describe such an algorithm and show an example of checker game between a real person and a virtual girl (Fig.2).

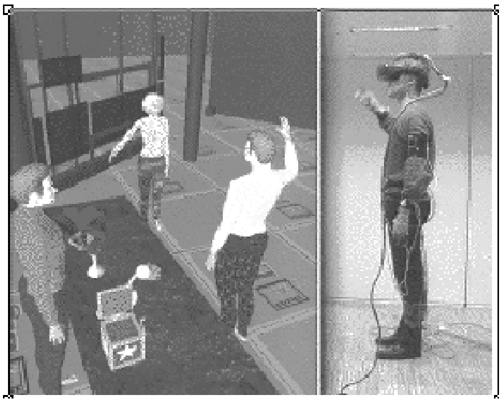

Fig.1. An avatar and two autonomous virtual humans

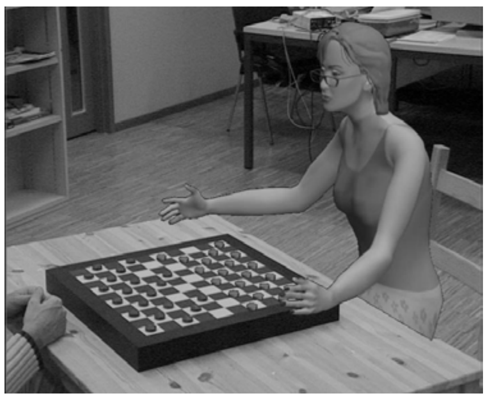

Fig.2. Interaction between a real person and a virtual girl

1 R. Boulic, S. Rezzonico, D. Thalmann, Multi-Finger Manipulation of Virtual Objects, Proc. ACM Symposium on Virtual Reality Software and Technology VRST'96, Hong-Kong, 1996, pp 67-74.

2. L. Emering, R. Boulic, D. Thalmann, Interacting with Virtual Humans through Body Actions, IEEE Computer Graphics and Applications, 1998 , Vol.18, No1, pp.8-11.

3 S. Balcisoy, M. Kallmann, R. Torre, P. Fua, D. Thalmann, Interaction Techniques with Virtual Humans in Mixed Environments, Proc. ISMR 01, Tokyo, Japan (extended version to be published in Presence, MIT Press)